A BRIEF LIST OF MATH GRE REVIEW CONCEPTS.

Bon-Soon Lin (Updated Thursday 13th August, 2020 at 10:07am)

This is by no mean comprehensive, but rather what comes to mind. Some of these are general guiding principles rather than rigorous statements, and some edge cases need to be considered more carefully.

1. Calculus I - Limits, functions, differential calculus.

1.1. Limits and limit laws for functions. Given a function $f(x)$ defined over a punctured open neighborhood of x_0. We say $\lim_{x\to x_0} f(x) = A$ if for every $\epsilon > 0$, there exists some $\delta > 0$ such that whenever $x \in S$ with $0 < |x - x_0| < \delta$, we have $|f(x) - f(x_0)| < \epsilon$.

Let $f(x), g(x)$ be functions such that $\lim_{x\to x_0} f(x) = A$ and $\lim_{x\to x_0} g(x) = B$, and c any constant. Then

(1) $\lim_{x\to x_0} f(x) + cg(x) = A + cB$.

(2) $\lim_{x\to x_0} f(x)g(x) = AB$.

(3) $\lim_{x\to x_0} f(x)/g(x) = A/B$ if $B \neq 0$.

For example $\lim_{x\to 0} \frac{(x^2+2x+1)\cos(3x)}{(x^3-2)} = \lim_{x\to 0} \frac{(x^2+2x+1)}{(x^3-2)} \cdot \lim_{x\to 0} \cos(3x) = -\frac{1}{2}$.

Also, the limit $\lim_{x\to 0} \sin(\frac{1}{x})$ does not exist, because near x, $\sin(\frac{1}{x})$ does not get arbitrarily close to any value.

Some more examples: $\lim_{x\to 0} \frac{x^2+2x+1}{x^2+3} = \frac{1}{3}$ while $\lim_{x\to\infty} \frac{x^2+2x+1}{x^2+3} = 1$.

1.2. Squeeze theorem. If we have $f(x) \leqslant g(x) \leqslant h(x)$ over some open I with $x_0 \in I$, and $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x) = L$ exists. Then $\lim_{x\to x_0} g(x) = L$.

This also holds for sequences: If $f(n) \leqslant g(n) \leqslant h(n)$, and $\lim_{n\to\infty} f(n) = \lim_{n\to\infty} h(n) = L$. Then $\lim_{n\to\infty} g(n) = L$.

A simple example: To find $\lim_{x\to 0} x^2 \cos\frac{1}{x^2}$ we note that $-1 \leqslant \cos\frac{1}{x^2} \leqslant 1$. So we have $-x^2 \leqslant x^2\cos\frac{1}{x^2} \leqslant x^2$, and since $\lim_{x\to 0} -x^2 = \lim_{x\to 0} x^2 = 0$, we have $\lim_{x\to 0} x^2\cos\frac{1}{x^2} = 0$ by squeeze theorem.

Be careful of signs when multiplying across inequalities, you may switch the direction of inequality!

1.3. Continuity. We say $f(x)$ is continuous at x_0 if for $\epsilon > 0$, there exists $\delta > 0$ such that whenever $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$.

A function $f(x)$ is continuous at x_0 if and only if whenever a sequence $x_n \to x_0$, we also have $f(x_n) \to f(x_0)$. In other words, a continuous function preserves limits.

For example, the function $f(x) = \begin{cases} x\sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$ is continuous at $x = 0$. To see this, take any sequence $x_n \to 0$, then note $-|x_n| \leqslant |x_n \sin\frac{1}{x_n}| \leqslant |x_n|$, hence $|x_n \sin\frac{1}{x_n}| \to 0$ by squeeze theorem, as both $|x_n|$ and $-|x_n|$ converge to 0. Hence whenever $x_n \to 0$, we have $f(x_n) \to 0$. Since $f(0) = 0$, we see that f preserves limits at $x = 0$, we conclude that f is continuous at $x = 0$.

Continuity at a point is also preserved by sums, product, and composition of continuous functions. So $f(x)$ above is also continuous at every point $x \neq 0$, as it is a product and compositions of continuous functions on $x \neq 0$.

1.4. Intermediate value theorem. If a function $f(x)$ is continuous on the interval $[a, b]$, and y is a value between $f(a)$ and $f(b)$, then there exists some $x \in [a, b]$ such that $f(x) = y$.

This is useful in showing existence of a root to an equation on some interval $[a, b]$, namely showing $f(x) = 0$ for some x in $[a, b]$, provided we know f is continuous on $[a, b]$ and $f(a)$ and $f(b)$ have opposite signs.

For example, the function $f(x) = x^3 + 2x - 1$ has at least one root, since for large enough positive x we have $f(x)$ is positive, while small enough negative x we have $f(x)$ is negative. If we are careful, we can see that since $f(0) < 0$, $f(1) > 0$, f has a root on the interval $(0, 1)$. We can further see this is the only root, as $f'(x) = 3x^2 + 2$, which means its derivative is always positive, hence strictly increasing and injective, and thus it has exactly one root between $(0, 1)$.

1.5. Differentiability. A function $f(x)$ is differentiable at x_0 if the limit of the difference quotient $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = D$ exists. In this case we write $f'(x_0) = D$. If a function is differentiable at x_0, then it is also continuous at x_0.

Another way to view differentiability of a function f at x_0 is to say it is well-approximated by a linear transformation: $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$. More specifically, $f(x)$ is differentiable at x_0 if there exists a linear function $L(x) = f(x_0) + D(x - x_0)$ for some D such that

$$\lim_{x\to x_0} \frac{f(x) - L(x)}{x - x_0} = 0,$$

in other words we have $f(x) = L(x) + o(x - x_0)$. Graphically, this $L(x)$ is the tangent line to the function $f(x)$ at x_0.

For example, the function $f(x) = \begin{cases} x\sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$ is continuous at $x = 0$, but not differentiable at $x = 0$. Because the difference quotient $\frac{x\sin\frac{1}{x} - 0}{x - 0} = \sin\frac{1}{x}$ does not have a limit at $x = 0$. However, the function $g(x) = \begin{cases} x^2\sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$

is differentiable at $x=0$, since the difference quotient $\frac{x^2\sin\frac{1}{x}-0}{x-0} = x\sin\frac{1}{x} \to 0$ as $x\to 0$. Hence $g'(0)=0$. And as g is differentiable at $x=0$, it is also continuous at $x=0$.

1.6. **Derivative rules, chain rule, and derivative of an inverse.** Let f,g be differentiable at x_0. Then

(1) $(f+cg)'(x_0) = f'(x_0)+cg'(x_0)$

(2) $(fg)'(x_0) = f'(x_0)g(x_0)+f(x_0)g'(x_0)$

(3) $(f/g)'(x_0) = \frac{g(x_0)f'(x_0)-f(x_0)g'(x_0)}{g(x_0)^2}$ if $g(x_0)\neq 0$

We also have chain rule: If f is differentiable at $g(x_0)$ and g is differentiable at x_0, then $(f\circ g)'(x_0) = f'(g(x_0))g'(x_0)$.

Using chain rule also gives the derivative of an inverse: Suppose f is continuously differentiable at x_0 (meaning differentiable and with continuous derivative) and $f'(x_0)\neq 0$, then f is locally invertible at x_0, with some local inverse f^{-1}. And as $f^{-1}(f(x)) = x$, we have $(f^{-1})'(f(x)) = \frac{1}{f'(x)}$ near x_0. This is the one-dimensional statement of inverse function theorem.

For example, to compute the derivative of $f(x)=\sin^{-1}(x)$, we note that $f(\sin(x)) = x$, so $f'(\sin(x))\cos(x) = 1$, and hence $f'(\sin(x)) = \frac{1}{\cos(x)}$. This gives $f'(x) = \frac{1}{\cos(\sin^{-1}(x))}$. Now drawing a right triangle with angle $\sin^{-1}(x)$ shows that $\cos(\sin^{-1}(x)) = \sqrt{1-x^2}$. Hence $f'(x) = \frac{1}{\sqrt{1-x^2}}$.

1.7. **Shape of functions.** If a function is differentiable, then positive (resp. negative) derivative on interval I indicates the function is strictly increasing (resp. decreasing). A point x where $f'(x)=0$ is called a critical point, and it may, however need not, be a local maximum or local minimum (for example $f(x)=x^3$ at $x=0$). Fermat's theorem of stationary points gives a necessary condition: If f is differentiable on open interval I where f has a a local extremum at $x_0\in I$, then $f'(x_0)=0$.

To check for global extrema of a function (not necessarily differentiable) on a closed interval, one compare the value of the function f at the critical points, at points of non-differentiability, and at the boundary points.

Note if f is differentiable on (a,b), and its derivative is strictly positive (or strictly negative) on (a,b), then f is injective on (a,b). Can extend this to $[a,b]$ if f is further continuous on $[a,b]$.

If $f:\mathbb{R}\to\mathbb{R}$ is differentiable on $\mathbb{R}$, and there exists some positive (resp. negative) M such that for all $x\in\mathbb{R}$, $f'(x)>M$ (resp. $f'(x)<M$). Then f is surjective onto $\mathbb{R}$.

We say f is concave up (resp. concave down) over the interval I if for every two points $x<y$ in I, the graph of f over $[x,y]$ is weakly below (resp. weakly above) the line segment connecting $f(x)$ and $f(y)$.

If f is differentiable, then f is concave up (resp. concave down) over the interval I if f lies above (resp. below) every tangents in I

If f is twice differentiable, then on an open interval I where $f'>0$ (resp. $f'<0$) is concave up (resp. concave down).

A point x is an inflection point if it is where f changes concavity. A necessary condition is: If f is twice differentiable and f has an inflection point at x_0, then $f''(x_0)=0$. The converse is not true as we can see in $f(x)=x^4$ at $x=0$.

A differentiable function f on interval I has inflection point at x_0 if and only if f' has an isolated extremum at x_0.

A useful method to sketch the shape of a function is to compute $f'(x)$ and make a table of intervals and values of x where $f'<0$, $f'=0$, and $f'>0$.

1.8. **L'Hospital rule.** Let $f(x)$ and $g(x)$ be such that the limit $\lim_{x\to x_0}\frac{f(x)}{g(x)}$ is of the indeterminate forms $0/0$, ∞/∞. If $\lim_{x\to x_0}\frac{f'(x)}{g'(x)} = L$ exists, then $\lim_{x\to x_0}\frac{f(x)}{g(x)} = L$.

This is useful in determining value of limits, provided conditions met. Another useful method of determining limits is by power series expansions.

Note we can often use repeated application of L'Hospital rule, provided that the final quotient of derivatives exist. Also, this can be applied to cases where we have indeterminate forms $0^0, 1^\infty, \infty^0, 0\cdot\infty, \infty-\infty$ by suitably transform them to the indeterminate forms $0/0, \infty/\infty$. However, sometimes direct applications of L'Hospital rule yield nowhere, and one would have to try something else (say by algebraic manipulations or induction).

An example, find $\lim_{x\to\infty}(e^x+x)^{2/x}$. This has indeterminate form ∞^0. We set the quantity as L and take logarithm on both sides, giving $\log L = \lim_{x\to\infty}\frac{2}{x}\log(e^x+x)$, which is now an indeterminate form ∞/∞. If we consider the ratio of the derivatives, we get $\lim_{x\to\infty}\frac{2(e^x+1)}{(e^x+x)}$, which is still ∞/∞. So again another ratio of derivatives we get $\lim_{x\to\infty}\frac{2e^x}{e^x} = 2$. Since this final limit exist, by L'Hospital rule twice we see that $\log L = 2$, and thus $L=e^2$.

1.9. **Mean value theorem.** Let $f(x)$ be continuous on $[a,b]$ and differentiable on (a,b), then there exists some $c\in(a,b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Geometrically, this means that there is some point $c\in(a,b)$ such that the tangent line at $(c,f(c))$ is parallel to the secant line connecting $(a,f(a))$ and $(b,f(b))$.

This can be useful in estimating growth of a function if we have a known bound on the derivative.

For example if f has derivative $f'(x)>2$ for all x, and $f(1)=3$, we must have $f(2)>5$. Indeed, if $f(2)\leqslant 5$, then by mean value theorem there exists some point $x\in(1,2)$ such that $f'(x) = \frac{f(2)-f(1)}{1}\leqslant 2$, a contradiction.

For another example, say we want to compute the limit $\lim_{x\to\infty}\sqrt{x+1}-\sqrt{x}$, which we could do analyze directly by dividing both sides by $\sqrt{x}$. Alternatively, note that $\lim_{x\to\infty}\sqrt{x+1}-\sqrt{x} = \lim_{x\to\infty}\frac{\sqrt{x+1}-\sqrt{x}}{x+1-1}$, and by mean value theorem there exists some point $c_x \in (x, x+1)$ such that $f'(c_x) = \frac{\sqrt{x+1}-\sqrt{x}}{x+1-1}$. But we know $f'(c_x) = \frac{1}{2\sqrt{c_x}}$. So $\lim_{x\to\infty} \frac{\sqrt{x+1}-\sqrt{x}}{x+1-1} = \lim_{x\to\infty} \frac{1}{2\sqrt{c_x}}$. Since $c_x \to \infty$ as $x \to \infty$ (as $c_x > x$), we see the limit has value 0.

1.10. **Implicit differentiation.** Suppose $y = y(x)$ is differentiable, and we have an implicit relation $F(x, y) = 0$, where F is differentiable with respect to both x and y. Then we can get an expression relating $y'(x)$ by noting $F(x, y) = F(x, y(x)) = 0$, and hence $\frac{d}{dx}F(x, y(x)) = 0$.

For example if $y = y(x)$ and we know that $\sin(xy) = x$, then $\cos(xy)(y+xy') = 1$, so $y' = \frac{1}{x\cos(xy)} - \frac{y}{x}$. In particular, at the point $(x, y) = (\frac{1}{2}, \frac{\pi}{3})$, we have $y'(\frac{1}{2}) = \frac{4}{\sqrt{3}} - \frac{2\pi}{3}$. As we see the derivative may still have an expression that depend on y, which we could try to use the original relation to eliminate y. However, if we are only interested in evaluations, then we just need to substitute the relevant values.

1.11. **Related rates.** Suppose we have differentiable functions $x(t)$ and $y(t)$, and they are subject to some relation $F(x, y) = 0$, where F is differentiable with respect to both x and y. Then we can get a relation between $x'(t)$ and $y'(t)$ by noting $F(x, y) = F(x(t), y(t)) = 0$, and differentiate with respect to t by chain rule. We can extend this to several functions of t.

For example a right triangle has side lengths $a(t), b(t), c(t)$ related by $a(t)^2 + b(t)^2 = c(t)^2$, then the rate of change of the lengths are related by $2a(t)a'(t) + 2b(t)b'(t) = 2c(t)c'(t)$.

1.12. **Optimization problems.** Let $F(x)$ be a function to be optimized, if F is differentiable, then we know local extrema occurs at some x where $F'(x) = 0$. If F is defined over an interval, then we need to check for the boundary values are well. This is the same problem as finding maximum and minimum of a function, after carefully setting up the problem.

For example, find the largest possible area of an isosceles triangle with total perimeter P. In this case, say the base is b, then the height is $\sqrt{\left(\frac{P-b}{2}\right)^2 - \left(\frac{b}{2}\right)^2}$, giving an area $A(b) = \frac{1}{2}b\sqrt{\left(\frac{P-b}{2}\right)^2 - \left(\frac{b}{2}\right)^2}$. We then attempt to find the maximum of A by considering points where $A'(b) = 0$ as well as $b = 0$, $b = P/2$ (boundary points). We will get $b = P/3$, which agrees with geometric intuition. Finally, to get the area, we put $b = P/3$ back and get $\frac{P^2}{12\sqrt{3}}$.

2. Calculus II - Integral calculus, sequences and series.

2.1. **Area under the curve; Riemann sum for continuous functions.**

2.2. **Fundamental theorem of calculus.**

2.3. **Volumes of revolution, surface of revolution, arclength of graphs.**

2.4. **Integration by parts.**

2.5. **Substitutions.**

2.6. **Integration of rational functions - partial fraction decomposition.**

2.7. **Sequences and series.** A real sequence (a_n) is said to converge to L if for every $\epsilon > 0$, there exists some N such that whenever $n > N$, we have $|a_n - L| < \epsilon$.

This means if L is the limit to the sequence (a_n), for every tolerance $\epsilon > 0$ you set, the terms of the sequence eventually all land inside the interval $(L-\epsilon, L+\epsilon)$.

If a sequence does not converge, we say it diverges. A special situation of divergent sequence is when the sequence tend to $+\infty$ or $-\infty$. To be more precise, we say $a_n \to +\infty$ (resp. $a_n \to -\infty$) if for each $M > 0$ (resp. $M < 0$), there exists N such that whenever $n > N$, we have $a_n > M$ (resp. $a_n < M$).

A useful thing to keep in mind: A sequence $a_n \to 0$ if and only if $|a_n| \to 0$.

Let $(a_n), (b_n)$ be two sequences and c a constant. Suppose $a_n \to A$ and $b_n \to B$, then

(1) $a_n + cb_n \to A + cB$

(2) $a_nb_n \to AB$

(3) $a_n/b_n \to A/B$ if $b_n \neq 0$ for all n and $B \neq 0$

An infinite series is of the form $\sum_{n=0}^{\infty} a_n$, and we say it converges if the partial sum sequence (s_n) converges, where $s_n = \sum_{k=0}^{n} a_k$.

A necessary but not sufficient test: If an infinite series $\sum a_n$ converges, then we must have its terms $a_n \to 0$. Converse is not true in general.

Some important series to keep in mind: $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if $p > 1$, and diverges if $0 < p \leqslant 1$. The geometric series $\sum_{n=1}^{\infty} r^n$ converges if and only if $|r| < 1$.

2.8. **Monotone bounded sequence theorem.** If (a_n) is an eventually monotonic sequence (either weakly increasing or weakly decreasing eventually), and (a_n) is bounded, then (a_n) converges.

2.9. **Integral test.**

2.10. **Comparison test.**

2.11. **Alternating series test.**

2.12. **Absolute convergence; root test and ratio test.** If $\sum_{n=1}^{\infty} |a_n|$ converges, then so does $\sum_{n=1}^{\infty} a_n$.

A series $\sum a_n$ is said to converge absolutely if $\sum |a_n|$ converges; it is conditionally convergent if $\sum a_n$ converges but $\sum |a_n|$ diverges.

2.13. **Power series.**

3. Calculus III - Multivariable calculus.

4. Elementary differential equations.

5. Matrix linear algebra.

5.1. **Cayley-Hamilton.** Given any square matrix A with $p(t)$ as its characteristic polynomial, then the matrix evalution $p(A) = O$ is the zero matrix.

5.2. **Real spectral theorem.** A real symmetric matrix A is always orthogonally diagonalizable, where the eigenspaces are mutually orthogonal to each other.

5.3. **Deciding eigenvalues of a matrix.** For a square matrix A, a number λ is an eigenvalue if and only if it is the root to the characteristic polynomial $p(t) = \det(A - tI)$. In other words, λ is an eigenvalue of A if and only if $A - \lambda I$ is not invertible. This can be seen easily when the row reduced matrix of $A - \lambda I$ has free variables. Furthermore, the number of free variables in $A - \lambda I$ is the dimension of the eigenspace $E_\lambda = \ker(A - \lambda)$. The dimension of E_λ can sometimes tell you if you need to look for more eigenvalues.

If one is able to find all n eigenvalues for a matrix A (namely the characteristic polynomial splits), then the sum of all n eigenvalues (including all multipicities) equals to the trace of the matrix A, while their product is $\det A$.

Some special cases. If a square matrix A has a common row sum, then that sum is an eigenvalue with eigenvector all 1's. If a square matrix has a common column sum, then as transposes have the same eigenvalues, that common column sum is also an eigenvalue. However, it is not clear what the eigenvector will be directly in this case.

If a matrix is upper or lower triangular, then the eigenvalues are precisely the entries on the main diagonal.

If a matrix has only strictly positive entries, then there exists a positive eigenvalue with a positive eigenvector. (Use Brouwer fixed point theorem to prove this.) Furthermore the largest positive eigenvalue for such a matrix with strictly positive entries will have geometric multiplicity one. (Peron-Frobenius)

5.4. **Symmetric 2×2 real matrices.** A good matrix to remember is one of the form $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$ where $a, b \in \mathbb{R}$. In this case the eigenvalues are $a+b$ and $a-b$, with eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ respectively.

Another matrix to remember is an $n \times n$ matrix of the form $\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$, where every entry is 1. From above we see that n is an eigenvalue, with eigenvector all 1's in entry. Since this is not invertible, 0 is an eigenvalue, with multiplicty $n-1$. This gives all the eigenvalues. The eigenspace to 0 is orthogonal to the vector of all 1's, by real spectral theorem.

5.5. **Companion matrices.** Given numbers $a_0, a_1, \ldots, a_{n-1}$, the matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-1} \end{pmatrix}$$

where the 1's are on the superdiagonal positions is called a **companion matrix**. Its characteristic polynomial, as well as its minimal polynomial, is precisely $t^n + a_{n-1}t^{n-1} + \cdots + a_0$.

5.6. **Vandermonde matrices.** Consider n distinct numbers $a_1, \ldots, a_n$, the matrix of the form

$$\begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix}$$

is called a **Vandermonde matrix**. Its determinant is given by $\prod_{i<j}(a_j - a_i)$. Though its eigenvalues and eigenvectors are not as straightforward. Note that as a_i are all distinct, this matrix will be invertible. This matrix is useful in Lagrange interpolation.

5.7. **Cayley-Hamilton.** Let A be an $n \times n$ matrix over any field K, and $p(t) = \det(A - tI)$ be its characteristic polynomial. Then we have $p(A) = O$, the zero matrix. Note p is of degree n.

5.8. **Minimal polynomial.** Let A be an $n \times n$ matrix, let $\mu_A(t)$ be the monic polynomial such that $\mu_A(A) = O$ of least degree. This polynomial exists because we know A satisfies at least one polynomial, namely its characteristic polynomial. Here μ_A has degree between 1 and n. The roots of μ_A gives the eigenvalues, and the power of each linear factor $(t - \lambda)$ of μ gives the size of the largest Jordan block of eigenvalue λ.

If f is any polynomial such that $f(A) = O$, then μ_A divides f.

The matrix A is diagonalizable if and only if μ_A is square-free.

6. Real analysis and basic topology.

6.1. **Compactness.** Let X be a topological space (say metric spaces like $\mathbb{R}$), a subset $K \subset X$ is said to be **compact** if and only if for every open covering $\{U_a\}_{a \in A}$ of K, there exists a finite subcovering, namely some finitely many $U_{a_1}, \ldots, U_{a_n}$ where $K \subset U_{a_1} \cup \cdots \cup U_{a_n}$.

For example, the set $A = \{1/n : n \in \mathbb{N}_{>0}\}$ is not compact in $\mathbb{R}$ with usual topology, as one can cover each point $1/n$ with an open set U_n that contains only $1/n$ from A. This covering for A cannot be refined to a finite subcovering. If we adjoin 0 to A, then the resulting set $A \cup \{0\}$ is compact.

In usual Euclidean metric topology on $\mathbb{R}^n$, we have **Heine-Borel theorem**, which a set $K \subset \mathbb{R}^n$ is compact if and only if K is closed and bounded. For instance $[0,1]$ will be compact with usual topology of $\mathbb{R}$.

The continuous image of a compact set is compact: If $f : X \to Y$ is continuous, and $K \subset X$ is compact, then $f(K)$ is also compact.

If $f : X \to Y$ is continuous on compact set X, then f is uniformly continuous on X, where X, Y are metric spaces.

6.2. **Composition of a function with a metric.** If $d : X \times X \to X$ is a metric on X, and $f : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function, concave on $[0, \infty)$ with $f(0) = 0$, then $f \circ d$ is again a metric. Note: concave means concave down.

So for example, if d is a metric on X, then so is $\sqrt{d}$, $\frac{d}{1+d}$, $d^{1/n}$ for any positive integer n, $\log(d + 1)$, $\arctan(d)$, etc.

One can weaken this further. If $f : [0, \infty) \to [0, \infty)$ is increasing with $f(x) = 0$ if and only if $x = 0$, and that f is subadditive, $f(x+y) \leqslant f(x)+f(y)$, then $f \circ d$ is a metric on X whenever d is a metric on X.

6.3. **Base of a topology.** Given a topological space (X, τ), we say a collection $B \subset \tau$ is a **base** if (1) $\bigcup_{U \in B} U = X$ and (2) for any $U_1, U_2 \in B$, and any $x \in U_1 \cap U_2$, there exists $U_3 \in B$ where $x \in U_3 \subset U_1 \cap U_2$. Note in particular, each $U \in B$ is open in the topological space X.

If B' is any collection of subsets of X that satisfies conditions (1) and (2), then B' generates a unique topology (the smallest topology containing B' on X).

7. Complex analysis.

7.1. **Analytic functions.** An open connected subset of $\mathbb{C}$ is said to be a **domain**. A complex-valued function $f : D \to \mathbb{C}$ on a domain D is said to be **complex differentiable** at the point $p \in D$ if the limit of the difference quotient

$$f'(z) = \lim_{h \to 0} \frac{f(z + h) - f(z)}{h}$$

exists. We say f is **analytic** at the point $p \in D$ if f is complex differentiable in an open neighborhood of p. And we say f is analytic on D if it is analytic at each point of D. We say f is **entire** if it is analytic on the entire complex plane $\mathbb{C}$.

Examples and intuition. An intuitive complex analytic funciton are "natural functions in the complex variable in z". In particular a polynomial $p(z)$ is entire, and power series $\sum_{n=0}^{\infty} a_n(z - a)^n$ is analytic on its open disc of convergence with positive radius.

Functions like $g(z) = |z|^2 = z\overline{z}$ or $g(z) = \overline{z}$ are not "natural functions of z" as they involve $\overline{z}$, and hence we will expect points of non-analyticity.

One key feature is that if f is analytic at a point $z \in D$, then it is infinitely (complex) differentiable at z. This "once differentiable implies infinitely differentiable" is not a feature of real functions on real domains.

7.2. **Cauchy-Riemann equations.** If a complex-valued function $f = u + iv$ on a complex domain D is analytic, then f satisfies a pair of equations called Cauchy-Riemann equation

$$u_x = v_y$$
$$u_y = -v_x$$

This restricts the class of functions that can be analytic (there is a rigid structure to an analytic function!)

7.3. **Harmonic functions.** If $f = u+iv$ is an analytic on a domain D, then the components are **harmonic**, namely $\Delta u = \Delta v = 0$, where $\Delta g = g_{xx} + g_{yy}$. Here the harmonic functions u and v such that $u + iv$ forms an analytic function are called **harmonic conjugates** of each other. And if u is harmonic on a simply connected domain D, then there exists a harmonic conjugate v on D such that $u + iv$ is an analytic function on D.

7.4. **Cauchy's theorem.** If a complex-valued function f is analytic on domain D, then for any closed loop integral γ whose enclosed region is contained in D we have

$$\int_\gamma f(z)dz = 0.$$

7.5. **Cauchy integral theorem.** Let f be a complex analytic function on domain D and on ∂D, then for any $z \in D$ we have

$$f(z) = \frac{1}{2\pi i} \oint_{\partial D} \frac{f(w)}{w - z} dw.$$

This generalizes to higher derivatives:

$$f^{(n)}(z) = \frac{n!}{2\pi i} \oint_{\partial D} \frac{f(w)}{(w - z)^n} dw.$$

This shows if f is analytic on domain D, then it is infinitely differentiable on D.

7.6. **Liouville theorem.** If f is entire (analytic on the entire complex plane) and bounded, then f is a constant function.

7.7. **Power series.** Given any complex sequence $a_n \in \mathbb{C}$, we can consider the power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$. If this power series has a positive radius of convergence $r > 0$, then this defines an analytic function f on the open disk $|z - z_0| < r$. This radius can be computed by Hadamard's formula:

$$r = \frac{1}{\limsup \sqrt[n]{|a_n|}}.$$

Conversely, given a function f that is analytic at some point z_0, then the Taylor series of f at z_0 given by

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n$$

will agree with f locally at z_0. This is not an expected behavior of a real function, even if it is infinitely differentiable! For example, the real function $f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$ is infinitely differentiable at $x = 0$, and $f^{(n)}(0) = 0$ for all n. This gives an identically zero Taylor series. However, f is not identically zero near $x = 0$.

Intuitively, this radius of convergence will be as large as possible before it is obstructed by a non-removable singularity. For example the function $f(z) = \frac{1}{1+z^2}$ has a Taylor series at $z = 0$ whose radius of convergence is 1. The reason is because f has singularities at $z = \pm i$. If we expand f at $z = 3$ instead, then the Taylor series at $z = 3$ will have radius of convergence of $\sqrt{10}$. And this is true of real function $f(x) = \frac{1}{1+x^2}$,its Taylor series at $x = 3$ will have radius $\sqrt{10}$ for this precise reason!

8. Algebra - Groups, rings, and fields.

8.1. **Ideals.** Given a ring $(R, +, \cdot)$, a **left ideal** (resp. **right ideal**) $I \subset R$ is a subgroup of the additive group $(R, +)$ such that $RI \subset I$ (resp. $IR \subset I$). A subset $I \subset R$ is said to be a **two-sided ideal** if it is both a left and a right ideal, and often we simply call two-sided ideals just **ideals**. In a commutative ring, all ideals are two-sided. For any ring R there is always the trivial ideals (0) and R.

An ideal I (two-sided) of a ring R induces a new ring structure called the **quotient ring** R/I that consists of all the cosets $r + I$, where we have addition and multiplication on R/I given by $(r+I)+(r'+I) = (r+r')+I$ and $(r+I)(r'+I)$.

9. Discrete math, geometry, and combinatoria.

9.1. **Trees.** A tree is a graph G that is connceted and has no cycles. In a tree of n vertices, there are always exactly $n-1$ edges. Conversely, a connected graph with n vertices and $n-1$ edges must be a tree.

9.2. **Weak Stirling approximation.** For n an integer, a quick estimation for $\log n!$ can be stablished as

$$n \log n - n < \log n! < n \log n,$$

which upon dividing $n \log n$, we see that $\log n! \sim n \log n$. (Note, $f(n) \sim g(n)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$.)

The inequality $\log n! < n \log n$ is because $n! < n^n$. And if we note $e^n = \sum_{k=0}^{\infty} n^k/k!$, we see that $e^n > n^n/n!$, or $n > n \log n - \log n!$.

9.3. **Stirling approximation.** The classic Stirling approximation gives

$$n! \sim \frac{n^n}{e^n}\sqrt{2\pi n}$$

as $n \to \infty$.

9.4. **Estimation of the central binomial coefficient $\binom{2n}{n}$.** Using Stirling approximation, we have

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} \sim \frac{4^n}{\sqrt{\pi n}}.$$

9.5. **A bound for $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k}$.** It is often needed to estimate the tail sum of the binomial coefficients, $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k}$, where $k \leqslant n/2$. We have a simple bound by geometric series:

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k} \leqslant \binom{n}{k} \frac{n-k+1}{n-2k+1}.$$

To see this, divide the sum by $\binom{n}{k}$, which is the largest term (as $k \leqslant n/2$), and we get

$$\begin{aligned} \frac{\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k}}{\binom{n}{k}} &= 1 + \frac{k}{n-k+1} + \frac{k(k-1)}{(n-k+1)(n-k+2)} + \cdots \\ &\leqslant 1 + \frac{k}{n-k+1} + \left(\frac{k}{n-k+1}\right)^2 + \cdots \\ &\leqslant \sum_{i=0}^{\infty} \left(\frac{k}{n-k+1}\right)^i = \frac{n-k+1}{n-2k+1}. \end{aligned}$$

9.6. **Number of trailing zeros in $n!$.** The number of trailing zeros of a number in decimal expansion is decided by the number of factors of 2 and 5. In $n!$, as factors of 2 shows up often enough, the number of zeros in $n!$ is decided by the factor of 5's. Note every multiple of 5 will contribute a factor of 5, while every multiple of 25 contribute an additional factor of 5, and so on. Hence we have the number of trailing zeros of $n!$ to be

$$\lfloor \frac{n}{5} \rfloor + \lfloor \frac{n}{5^2} \rfloor + \lfloor \frac{n}{5^3} \rfloor + \lfloor \frac{n}{5^4} \rfloor + \cdots$$

This is a finite sum as eventually the floors will be zero. If we disregard the floor function, then we can roughly estiamte the number of trailing zeros of $n!$ by geometric series $n\left(\frac{1}{5}+\frac{1}{25}+\frac{1}{125}+\cdots\right)=n/4$.

An observation one can make is that not every positive integer k shows up as the number of trailing zeros of some $n!$. For example 24! factorial will have four zeros, while 25! will have six zeros, as 25 contributes two factors of 5. But if an integer k does show up as the number of trailing zeros of some number $n!$, then it shows up exactly 5 times (if k shows up first as the number of trailing zeros of some $n!$, then the numbers $(n+1)!$ to $(n+4)!$ will all have k many trailing zeros, as no new factors of 5 are introduced. But $(n+5)!$ will introduce at least one new factor of 5).

10. Probability and statistics.

10.1. **68-95-99.7 rule.** Let $X \sim N(0,1)$ be normally distribued. Then we have approximately $P(-1<X<1)=0.68$, $P(-2<X<2)=0.95$, and $P(-3<X<3)=0.997$.

10.2. **Estimating binomial distribution with normal distribution.** Suppose $X \sim \text{binomial}(n,p)$. If we have $np, nq \geqslant 5$, say, then it might be suitable to approximate X with a normal distribution $N(\mu=np, \sigma=\sqrt{npq})$. Here $q=1-p$. In this case, one can convert X into a Z-score by $Z=\frac{X-np}{\sqrt{npq}}$, where we approximate Z with standard normal $N(0,1)$.

11. Basic logic and set theory.

References.

Materials adopted and expanded from various sources.